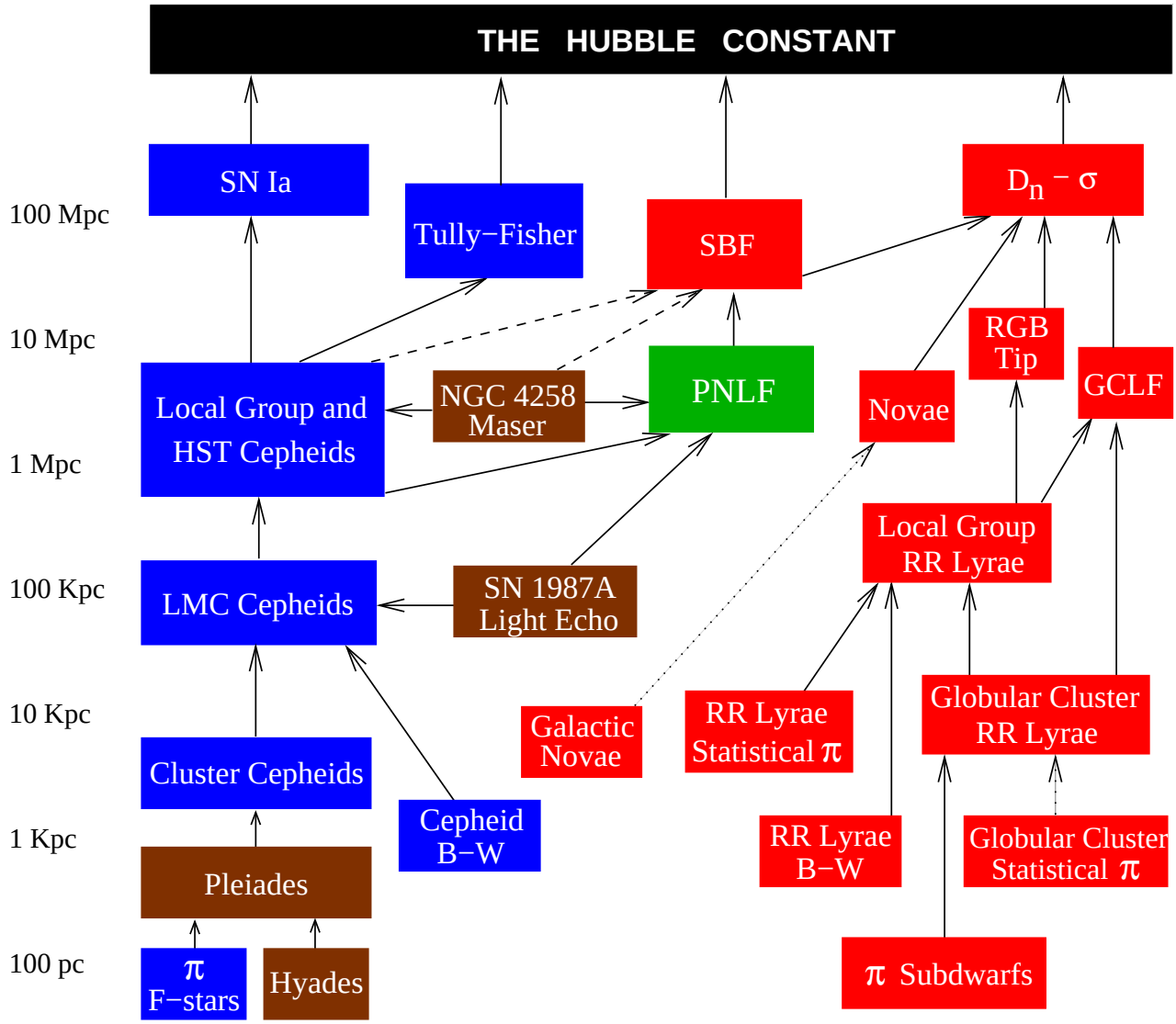
The Distance Ladder

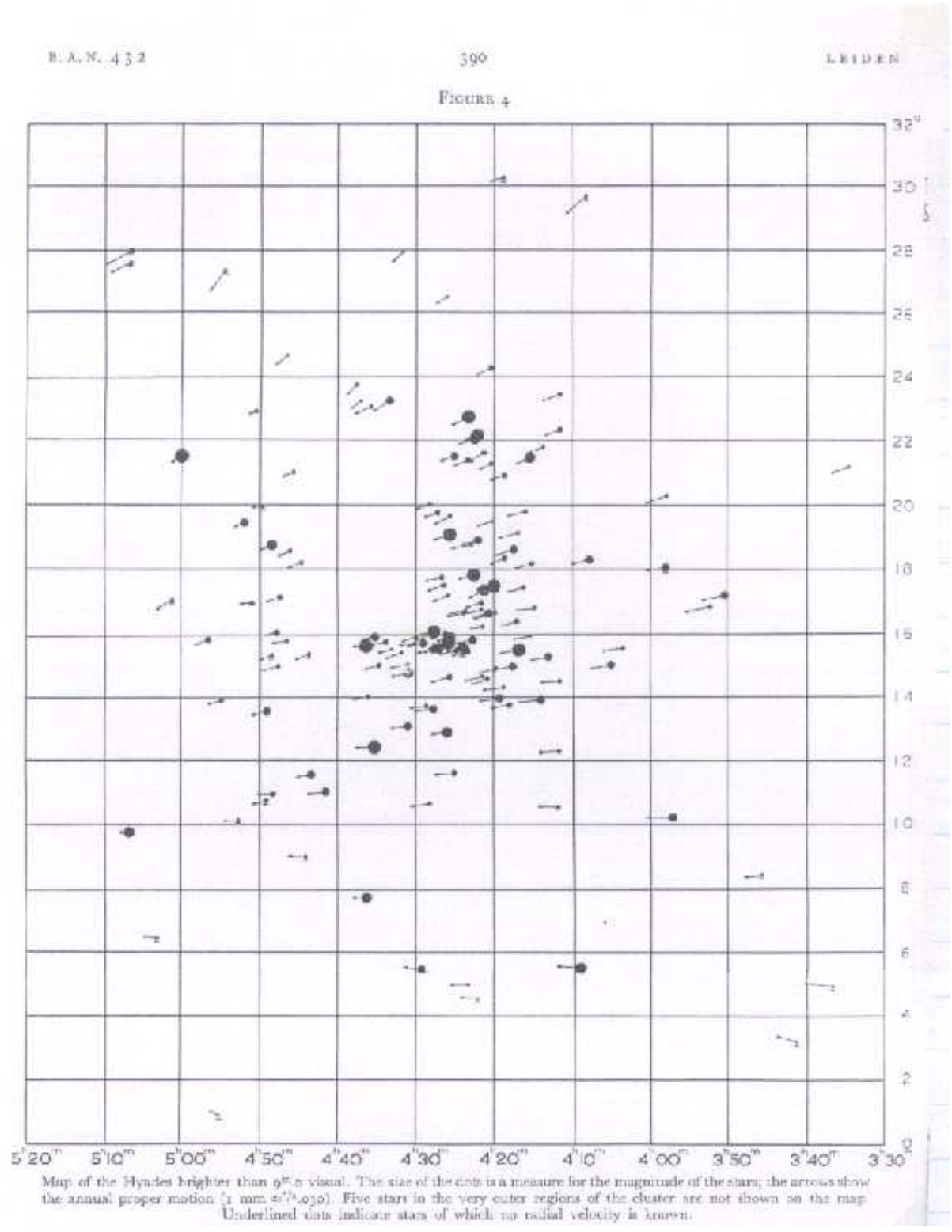
There are two classes of distance measurements: those which produce distances that are independent of other methods, and those that are part of the *distance ladder*. Most of the measurements you are familiar with fall in the latter category. Below is a schematic of the distancer ladder, with blue representing methods that work primarily (or exclusively) in Pop I systems, red for methods that work in old populations, and brown for geometric techniques that lie outside the ladder. The PNLF techniques works in all (large) galaxies, and is in green.



The Base of the Distance Ladder

At the base of the distance ladder are two geometric techniques: parallax (denoted by the symbol for parallax angle, π), and the moving cluster method (which is most applicable to the Hyades). The former relies on the earth's movement about the sun to provide a surveyor's baseline; the latter uses the space motions of co-moving groups of stars. (The angle between the cluster's convergent point and its location gives you ratio of radial velocity to tranverse velocity; this ratio then translates into a distance via the proper motion.) Unfortunately, these techniques are only good for objects closer than a few hundred parsecs. Although a couple of hundred thousand stars are in range, most are normal solar-metallicity G,K, or M main sequence stars. In fact, there are very few "interesting" targets within reach of the geomtrical techniques. Thus, to measure distance, we must build a ladder.

Type of Object	Distance (pc)	Number within Distance
Open Clusters	200	5
Metal-Poor Main Sequence	200	$\gtrsim 10$
Cepheid Variables	500	8
RR Lyrae Variables	500	1
Horizontal Branch Stars	600	$\gtrsim 7$
Red Supergiants	500	0
Blue Supergiants	500	1
Planetary Nebulae	500	10
Cataclysmic Varibles	350	29
White Dwarfs	500	many!



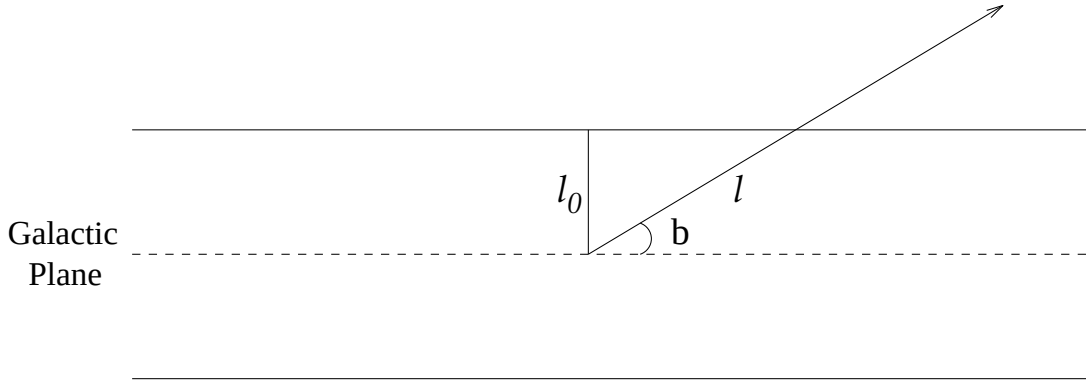
Proper motion vectors for the Hyades stars from the classic paper by van Bureren in Bull. Astron. Inst. Netherlands XI (1952). Note that the vectors appear to converge at a point about 20 degrees from the center of the cluster.

Interstellar Extinction

[Trumpler 1930, *Lick Obs. Pub.* 420]
[Sandage 1973, *Ap.J.,* **183,** 711]
[Burstein & Heiles 1982, *A.J.,* **87,** 1165]
[Burstein & Heiles 1984, *Ap.J. Suppl.,* **54,** 33]
[Schlegel, Finkbeiner, & Davis 1998, *Ap.J.,* **500,** 525]

Any standard candle measurement must take interstellar extinction into account. Extinction is present, even when the dust is not obvious. This was first proved by Trumpler in 1930, when he showed that the mean (linear) size for open clusters seemed to increase with distance. (In other words, the distances to these clusters were being overestimated, due to dust.) Since then, there have been a number of attempts to parameterize the dust distribution in the Galaxy.

One approach by Sandage was to examine the BVR colors of bright galaxies as a function of galactic latitude. He then assumed that the Galactic dust is distributed in a fashion similar to that of a uniform disk of thickness $2\ell_0$. In this case



$$I = I_0 e^{-\kappa \ell} = I_0 e^{-\kappa \ell_0 \csc b} \tag{17.01}$$

If we change this into magnitudes, then

$$-2.5 \log I = -2.5 \log I_0 + 2.5 (\log e) \kappa \, \ell_0 \csc b \tag{17.02}$$

If we let $A = \kappa \ell_0 (\log e) = 1.086\, \kappa\, \ell_0$, then

$$\Delta m = A \csc b \tag{17.03}$$

Sandage then modified this slightly

$$\begin{aligned} A_B &= 0.132(\csc b - 1) \quad \text{for} \quad |b| \leq 50^\circ \\ &= 0 \quad \text{for} \quad |b| > 50^\circ \end{aligned} \tag{17.04}$$

Of course, by using a sample of bright galaxies, Sandage was already introducing a selection effect: the larger the extinction, the less likely the galaxy is to be in the sample

Another approach which was taken by de Vaucouleurs, de Vaucouleurs & Corwin in the RC2 catalog, was to fit galaxy counts and radio data over the entire sky (using the underlying assumption that dust follows the gas). The result was a much more complicated function (too long to reproduce here). Today, most estimates of foreground Galactic extinction comes from two places.

The first is an quantitative analysis of galaxy counts and H I emission over the entire sky. (Again, this assumes a constant dust-to-gas ratio.) These values are reproduced as a series of contour diagrams in Burstein & Heiles (1982) and tabulated for NGC and UGC galaxies in Burstein & Heiles (1984). The second is based on COBE measurements of thermal emission from the dust, and is described in Schlegel, Finkbeiner & Davis (1998).

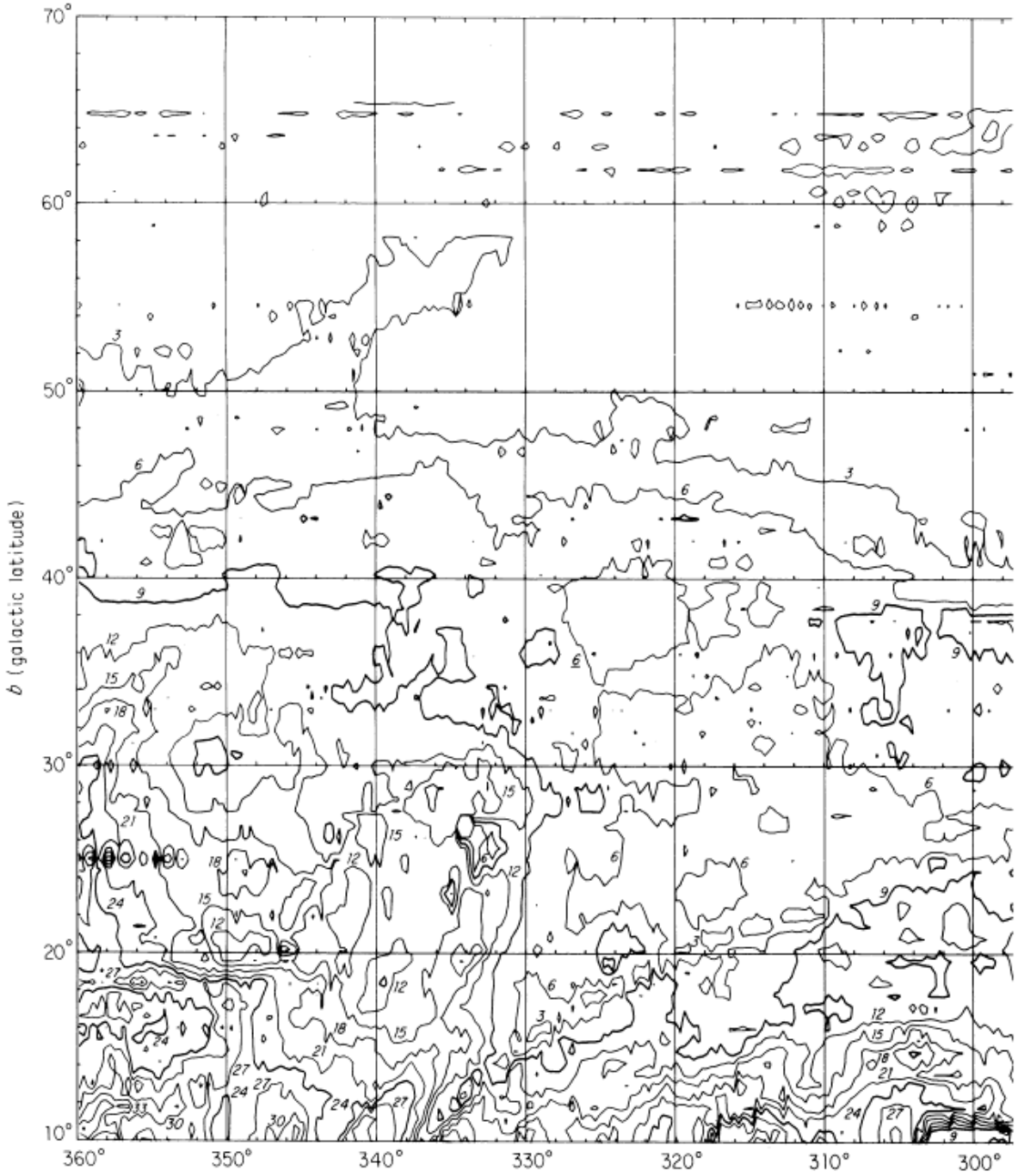


FIG. 4(a). A map of reddening in the galaxy for $10° < b < 65°$ and all galactic longitudes. These maps were generated by the H I/GC method as described in the text. Regions of the map corresponding to declinations south of $-23°$ are less accurate than the rest of the map. The contour levels are labeled in units of 0.01 mag in $E(B-V)$ in intervals of 0.03 mag. The contours start at 0.00, and the intervals corresponding to 0.09, 0.24, 0.39, 0.54, and 0.69 mag of $E(B-V)$ are emphasized in heavy ink. Not all contours are labeled in this map owing to excessive crowding in parts. Closed contours with an underlined label indicate reddenings that are decreasing interior to the contour; otherwise contours indicate increasing reddening interior to the contour. Most geometric figures in this map are artifacts of the way the mapping program handles regions without data.

Wavelength Dependence of Extinction

[Savage & Mathis 1979, *A.R.A.A.*, **17,** 73]
[Krelowski & Papaj 1993, *Pub. A.S.P.*, **105,** 1209]
[Bouchet *et al.* 1985, *Astr. Ap.*, **149,** 330]
[Cardelli, Clayton, & Mathis 1989, *Ap.J.*, **345,** 245]

The extinction due to dust varies with wavelength. In general, the shorter (bluer) the wavelength, the greater the extinction. Hence, the term "reddening". The precise wavelength dependence of extinction does vary from location to location in the Milky Way, and metal-poor galaxies appear to have different extinction laws. However, in the optical, the logarithmic extinction is approximately $A_\lambda \propto 1/\lambda$. Note that this is *not* Raleigh scattering, which goes as $1/\lambda^4$.

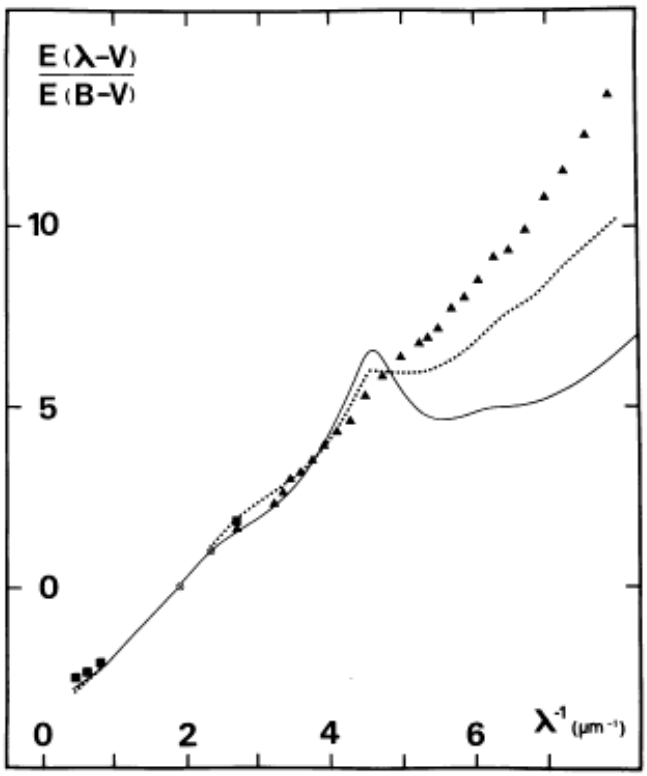


Fig. 3. The SMC extinction curve from 2.2 μm to 130 nm resulting from the present work (squares) and from Prévot et al. (1984) (triangles). The heavy line represents the galactic law [Savage and Mathis (1979), plus Koornneef (1983b) for the infrared] and the dotted line represents the LMC curve [mean of Nandy et al. (1981) and Koornneef and Code (1981) for the ultraviolet, and Koornneef (1982) for the infrared]. Encircled crosses at *B* and *V* wavelengths result from normalization

There are three ways of quoting the amount of extinction to an object. The first is to quote the total extinction (in magnitudes), in some band, usually B (*i.e.,* A_B). The second, which is used by people who study emission line objects, is to give c, the logarithmic amount of extinction at Hβ. But the most common way is to quote the differential extinction between two bandpasses, *i.e.,*

$$\begin{aligned} E(B-V) &= \Delta B - \Delta V \\ &= (B_{obs} - B_0) - (V_{obs} - V_0) \\ &= (B-V)_{obs} - (B-V)_0 \end{aligned} \tag{17.05}$$

Here, the subscript zero refers to the *intrinsic* magnitude of the object, and the subscript *obs* denotes the observed quantity. The extinction at any wavelength is then be calculated through

$$A_\lambda = R_\lambda E(B-V) \tag{17.06}$$

where R_λ is the ratio of the total to differential extinction. Some values of R are $R_B = 4$, $R_V = 3$, $R_R = 2.3$, and $R_I = 1.5$.

The precise extinction properties of dust depends on the type of dust involved. For instance, dust can be silicon-rich (*i.e.,* silicates), or carbon-rich. Also, observations in other galaxies suggest that the overall metallicity of the galaxy can affect dust properties. This is most pronouced in the ultraviolet, where low-metallicity galaxies have a rapidly increasing extinction curve. To account for this, Cardelli, Clayton & Mathis (1989) have parameterized extinction via A_V, the ratio of total V extinction to differential extinction. Different values of A_V imply different extinction laws.

The Extragalactic Calibrators – Pulsating Stars

Pulsating stars, such as RR Lyr stars and Cepheids, are extremely important for distance scale measurements. To understand why, let's put aside the question of *why* stars pulsate (that's a subject for a stellar structure class), and ask the question, *if* a star pulsates, what will be the period of its pulsation?

Image a star that is unstable to pulsations. We'll assume that about 1/2 of the star's time is spent in the expansion phase, and the other half in the contraction phase. We'll also make the assumption that the star's change in radius is some percentage of the star's average radius, *i.e.,* $\delta R \propto R$. Now, suppose the star has just reached maximum. How fast can gravity restore the star to its normal state? In the case where the acceleration due to gravity is roughly constant,

$$\delta R = \frac{1}{2} g t^2 \tag{17.07}$$

The time, t, is a fraction of the star's period of pulsation, so $P \propto t$. Furthermore, the gravitational acceleration is simply $G\mathcal{M}/R^2$. So

$$P^2 \propto \left(\frac{R^3}{\mathcal{M}} \right) \tag{17.08}$$

(Kelper's law), or

$$P \propto \langle \rho \rangle^{-1/2} \tag{17.09}$$

where $\langle \rho \rangle$ is the average density of the star. This simple equation is a remarkably good approximation for the period of all radially pulsating stars, and even some non-radial pulsators.

You can also get at equation (17.09) in a different way. Consider that if a star's interior undergoes a perturbation, that event will propagate through the star at the local sound speed, $v_s = \sqrt{\gamma p/\rho}$, where p is the pressure. Therefore, the time it takes the star to adjust is

$$t = P \propto R/v_s \propto R\left(\frac{\rho}{p}\right)^{1/2} \tag{17.10}$$

Now, if the star is in quasi-hydrostatic equilibrium

$$\frac{dp}{dR} = -g\rho = -\frac{G\mathcal{M}}{R^2}\rho \tag{17.11}$$

which, very roughly, integrates to

$$p \propto \frac{\mathcal{M}\rho}{R} \tag{17.12}$$

If you substitute this into (17.10), you again get

$$P \propto R\rho^{1/2}\left(\frac{R}{\mathcal{M}\rho}\right)^{1/2} \propto \left(\frac{R^3}{\mathcal{M}}\right)^{1/2} \propto \langle\rho\rangle^{-1/2} \tag{17.13}$$

So whether you approach the problem via atmospheric free-fall, Kepler's laws, or wave propagation, you get the same result.

Let's now put the equation for pulsation back in the form

$$P^2 \propto \left(\frac{R^3}{\mathcal{M}}\right) \tag{17.09}$$

We know from the Stefan-Boltzmann law that

$$\mathcal{L} = 4\pi R^2 \sigma T_{eff}^4 \tag{17.14}$$

so we can get rid of R in the equation. Thus,

$$P \propto \frac{\mathcal{L}^{3/4}}{T_{eff}^3 \mathcal{M}^{1/2}} \tag{17.15}$$

or

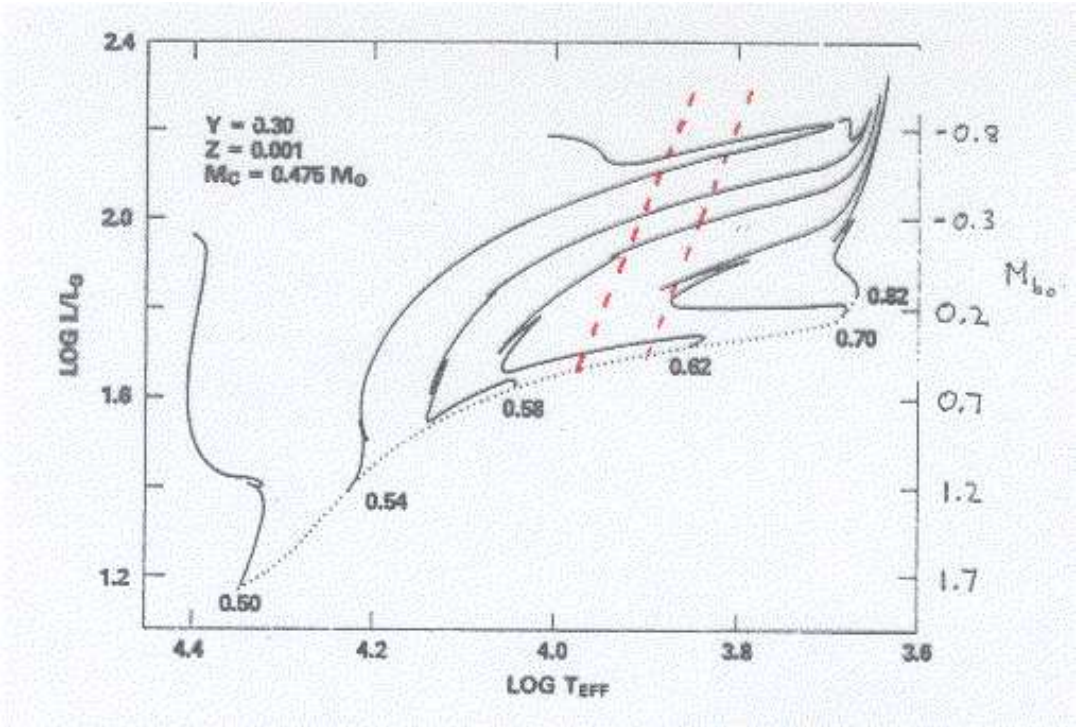
$$\log P = 0.75 \log \mathcal{L} - 3 \log T_{eff} - 0.5 \log \mathcal{M} + Q \tag{17.16}$$

where $\mathcal{L}$ is the mean luminosity of the star, and Q is the constant of proportionality. Now, consider what this equation says: every star has a unique natural period for pulsation which depends on its luminosity (and temperature).

Another way to look at the problem is to consider the HR diagram, which is $\log \mathcal{L}$ *vs.* $\log T_{eff}$. In many respects, you can think of this diagram as the projection of a 3-dimensional diagram, with axes of $\log \mathcal{L}$, $\log T_{eff}$. and $\mathcal{M}$. However, there is nothing special about these variables; for instance, you can easily remove $\log \mathcal{L}$ (or $\log T_{eff}$) from the diagram, and replace it with R, through the Stefan-Boltzmann law. Or, equivalently, you can use (17.16) and replace any of the three variables with P. So, just as there are rules for the locations of stars in the HR diagram, there are rules for the positions of stars in $\log \mathcal{L}$-$\log T_{eff}$-$\log P$ space.

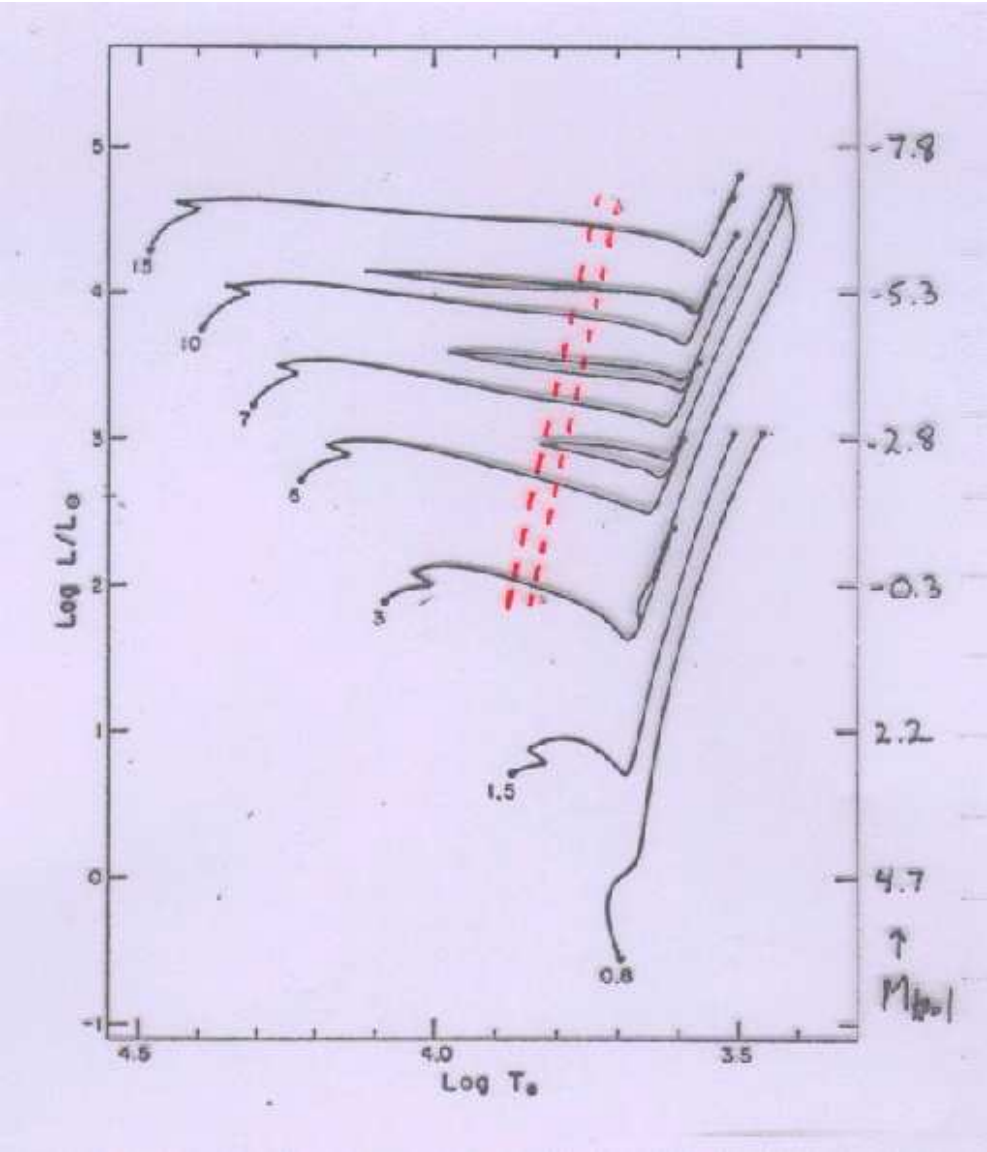
RR Lyrae and Cepheids

The two most commonly used standard candles among pulsating stars are RR Lyr variables and Cepheids. RR Lyr stars are old, primarily metal-poor objects that pulsate with periods $8 < P < 20$ hours, and amplitudes of $\lesssim 1$ mag. These stars began life on the main sequence with masses $\sim 1\mathcal{M}_\odot$; they have subsequently evolved up the giant branch, lost some mass, ignited helium in the core, and settled on the *horizontal branch,* where they have a luminosity of $\sim 50\mathcal{L}_\odot$, and a mass of $\mathcal{M} \sim 0.6\mathcal{M}_\odot$. Because horizontal branch stars all have roughly the same luminosity, RR Lyr stars make excellent standard candles. (Note, however, that *evolved* horizontal branch stars should be more luminous than zero-age HB stars. Plus, there's no guarantee that RR Lyr stars of different metallicity will have the same luminosity.)



The stellar evolutionary tracks of horizontal branch stars, and the RR Lyr instability strip. The tracks are different for other core masses and metallicities. As with the main-sequence, horizontal branch stars spend most of their life near their zero-age horizontal branch position.

Cepheid variables are massive stars that have evolved up the giant branch, ignited helium, decreased their luminosity (slightly), and entered into an intermediate phase before returning to the asymptotic giant branch. During this time, they make a loop in the HR diagram which passes through the instability strip. Because Cepheids are Pop I objects, they are most often found in regions of active star formation, where the affects of interstellar dust are significant. Cepheid pulsation periods can be anywhere from a few days to a couple of months.



The evolutionary tracks of high-mass stars and the instability strip for Cepheid variables.

The period of a Cepheid variable is a sensitive function of its absolute magnitude. To see this, consider that, like main-sequence stars, a mass-luminosity law exists for massive stars in the post-giant branch phase of evolution. So

$$\mathcal{L} \propto \mathcal{M}^{\alpha} \tag{17.17}$$

where $\alpha \sim 4$. When you combine that with the pulsation law

$$P \propto \langle \rho \rangle^{-1/2} \propto R^{3/2} \mathcal{M}^{-1/2} \tag{17.09}$$

and the Stefan-Boltzmann law

$$\mathcal{L} \propto R^2 T_{eff}^4 \tag{17.15}$$

you get

$$P \propto \mathcal{L}^{\frac{3\alpha-2}{4\alpha}} T_{eff}^{-3} \tag{17.18}$$

or (with $\alpha = 4$),

$$\log \mathcal{L} = 1.6 \log P + 4.8 \log T_{eff} \tag{17.19}$$

Note that there is a strong temperature dependence, as well as a luminosity dependence. However, if you make the assumption that the instability strip is very narrow, then a direct relation can be found between $\log \mathcal{L}$ and $\log T_{eff}$,

$$\log \mathcal{L} \approx 20 \log T_{eff} + C \tag{17.20}$$

The result is a simple $\log \mathcal{L}$-$\log P$ relation.

Though elegant, the usefulness of the previous analysis is quite limited. While it does show that the calibration of Cepheids as standard candles requires two numbers, a slope and an intercept, it also assumes that $\log \mathcal{L}$ is a measureable quantity. It is not – what is observed is flux in a particular bandpass.

Calibrating RR Lyr and Cepheid Variables

In order to use RR Lyrae and Cepheids, the absolute luminosities of these objects must be calibrated. In the case of RR Lyr stars, this means determining the star's absolute magnitude (and perhaps how the absolute magnitude changes with metallicity). For Cepheids, one must determine both zero point and slope of the period-luminosity relation.

Numerous methods are used to estimate the absolute magnitudes of pulsating stars, including statistical parallax (whereby the Sun's motion through space provides the baseline), main-sequence fitting of clusters which contain the stars, statistical parallax of globular cluster stars (where one observes the radial velocities of a set of stars in a globular cluster, and also measures their transverse proper motions), and pulsation theory. There is one especially notable technique: the Baade-Wesselink Method.

The Baade-Wesselink Method

Baade-Wesselink works for any pulsating star (or even exploding stars, such as supernovae). Consider a pusating star at minimum, with measured temperature, T_1, and observed flux, f_1. If the star's radius at minimum is R_1, then

$$f_1 = \frac{4\pi R_1^2 \sigma T_1^4}{4\pi D^2} \tag{17.21}$$

where D is the star's distance. Later on at maximum, the star has observed flux, f_2, a temperature T_2, and radius R_2, so

$$f_2 = \frac{4\pi R_2^2 \sigma T_2^4}{4\pi D^2} \tag{17.22}$$

The temperatures and fluxes are both observable, so there are 3 unknowns (R_1, R_2, and D) and two equations.

Now suppose you observe a star spectroscopically during its pulse from R_1 to R_2. During this time, the star's atmosphere has expanded at a velocity $v(t)$ from R_1 at time t_1 to R_2 at time t_2. So

$$R_2 = R_1 + \Delta R = R_1 + \int_{t_1}^{t_2} v(t)\, dt \tag{17.23}$$

If you measure $v(t)$ throughout the pulse, you have a measure of ΔR. This gives you three equations and three unknowns, and let's you solve for D. (This is left as an exercise to the student with nothing better to do.)

In practice, the method is a bit more complicated. For example, how does one translate color to temperature? Can conversions which are derived from normal stars be applied to stars with dynamic atmospheres? In addition, how does the observed radial

velocity of a stellar envelope relate to the true expansion velocity? The two are not equal. When you observe a star, you are getting flux over its entire surface, but only the very center of the star has a motion that is completely radial. The projection factor depends on a number of factors. When one moves from the center of a star to its limb,

- v_{obs} decreases as $\cos\theta$
- the unit surface area increases as $\cos\theta$
- limb darkening decreases the amount of flux received by the observer.

From stellar atmospheres, the amount of limb darkening can be estimated via the Eddington approximation

$$I(\theta) = \frac{3}{5}\cos\theta + \frac{2}{5} \tag{17.24}$$

Putting this together yields

$$\begin{aligned} v_{obs} &= \frac{\int\int (v\cos\theta) I(\theta)\cos\theta \cdot d\omega}{\int\int I(\theta)\cos\theta d\omega} \\ &= v\ \frac{\int_0^{\pi/2}\left(\frac{3}{5}\cos\theta + \frac{2}{5}\right)\cos^2\theta\sin\theta d\theta}{\int_0^{\pi/2}\left(\frac{3}{5}\cos\theta + \frac{2}{5}\right)\cos\theta\sin\theta d\theta} \\ &= v\left(\frac{17}{24}\right) = 1.41v \end{aligned} \tag{17.25}$$

[Gettings 1935, *MNRAS,* **95,** 139]

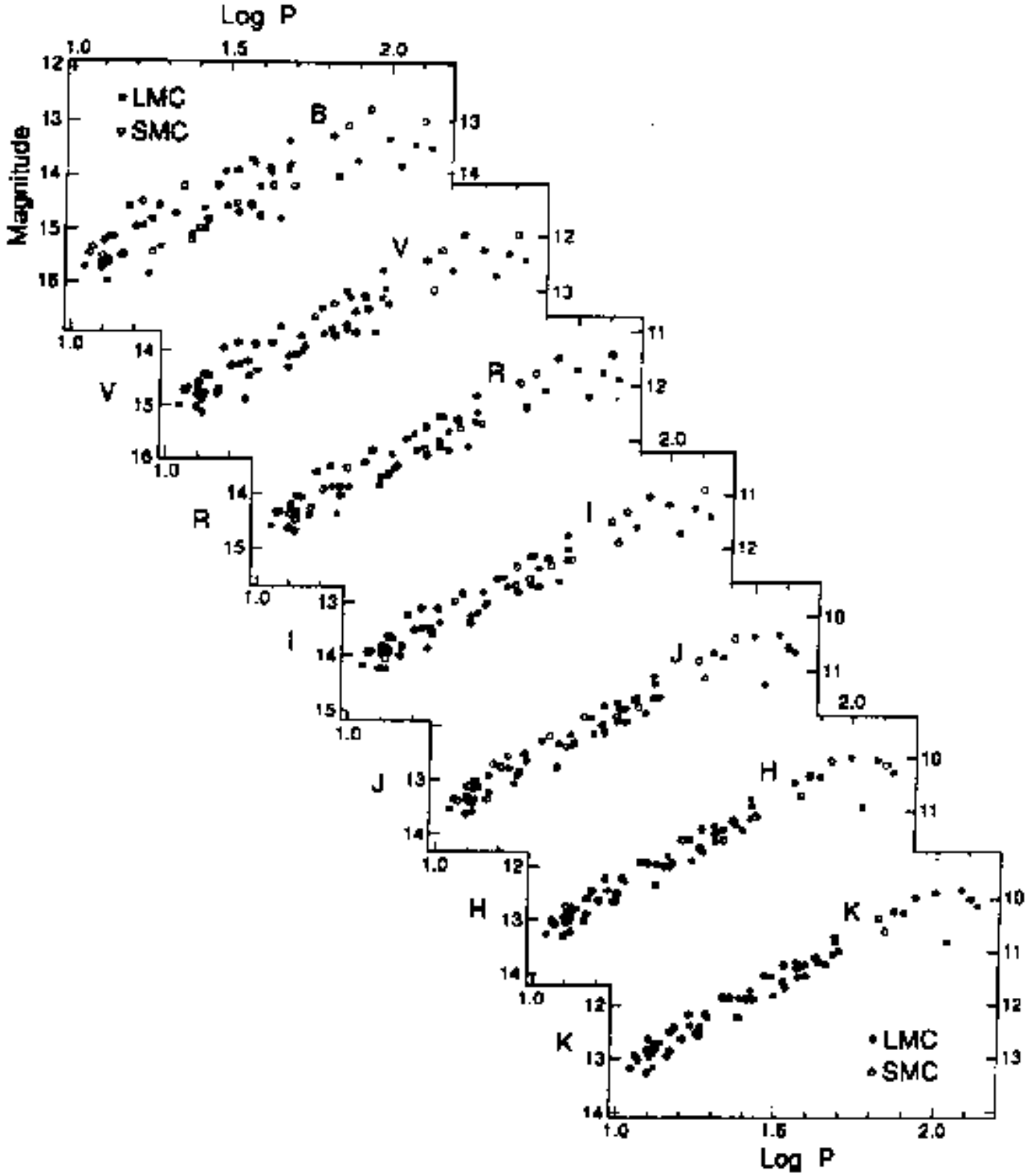
In practice, the calculation of the projection factor needs a model atmosphere analysis that involves the strength of the line being measured, the temperature of the star, and how one centroids the absorption line.

The Slope of the Period-Luminosity Relation

[Madore & Freedman 1991, *Pub. A.S.P.,* **103,** 933]
[Freedman & Madore 1990, *Ap.J.,* **365,** 186]
[Freedman *et al.* 2001, *Ap.J.,* **553,** 47]

With work, the Baade-Wesselink method and other techniques can be used to obtain distances to individual stars. But in the Milky Way, there are only a handful (~ 10) Cepheids whose distances (and therefore absolute magnitudes) with easily obtainable distances. (Remember, Cepheids are Pop I objects, so they are usually located in star-forming regions in the plane of the galaxy, surrounded by patchy dust.) Thus, Galactic Cepheids are not ideal for determining the slope of the Period-Luminosity Relation.

The best way to fix the slope is to observe a large sample of Cepheids at the same distance. Traditionally, this has been done in the Large Magellanic Cloud, our nearest (non dwarf) galaxy.


Log P
Magnitude
LMC
SMC
B
V
V
R
R
I
I
J
J
H
H
K
K
LMC
SMC
Log P

Note the increased scatter in the P-L relation at bluer wavelengths. Part of this is due to the larger effect of extinction. However, part of the scatter is intrinsic. The period of a Cepheid depends on both the star's luminosity and its temperature, but because the instability strip is narrow, the latter variable is usually neglected. However, consider the case of two stars on opposite sides of the strip. When viewed in the optical, the small difference in temperature will propagate into a rather large difference in flux within the bandpass (since you're near the peak of the blackbody curve). This introduces scatter into the period-luminosity relation. On the other hand, if the two stars are viewed in the infrared, the flux within the bandpass doesn't depend that much on temperature, hence the scatter is smaller.

By observing Cepheids at multiple wavelengths, it is possible to remove all the effects of reddening. Consider a sample of Cepheids, extinguished by an amount of dust parameterized by $E(B-V)$. The period-luminosity relation says that

$$\langle M_x \rangle = a \log P + b \tag{17.26}$$

where x is the wavelength of the observation. When you observe Cepheids in another galaxy, you will observe their periods and their apparent magnitudes, $\langle m_x \rangle$. You therefore derive the apparent distance modulus, μ_x, which is related to the true distance modulus, μ_0, by

$$\mu_x = \langle m_x \rangle - \langle M_x \rangle = \mu_0 + R_x E(B-V) \tag{17.27}$$

In other words, you have one equation, but two unknowns, μ_0 and $E(B-V)$. However, if you observe at more than one wavelength, you have more than one equation (but still the same two unknowns). So you can solve for *both* the reddening and the distance simultaneously.

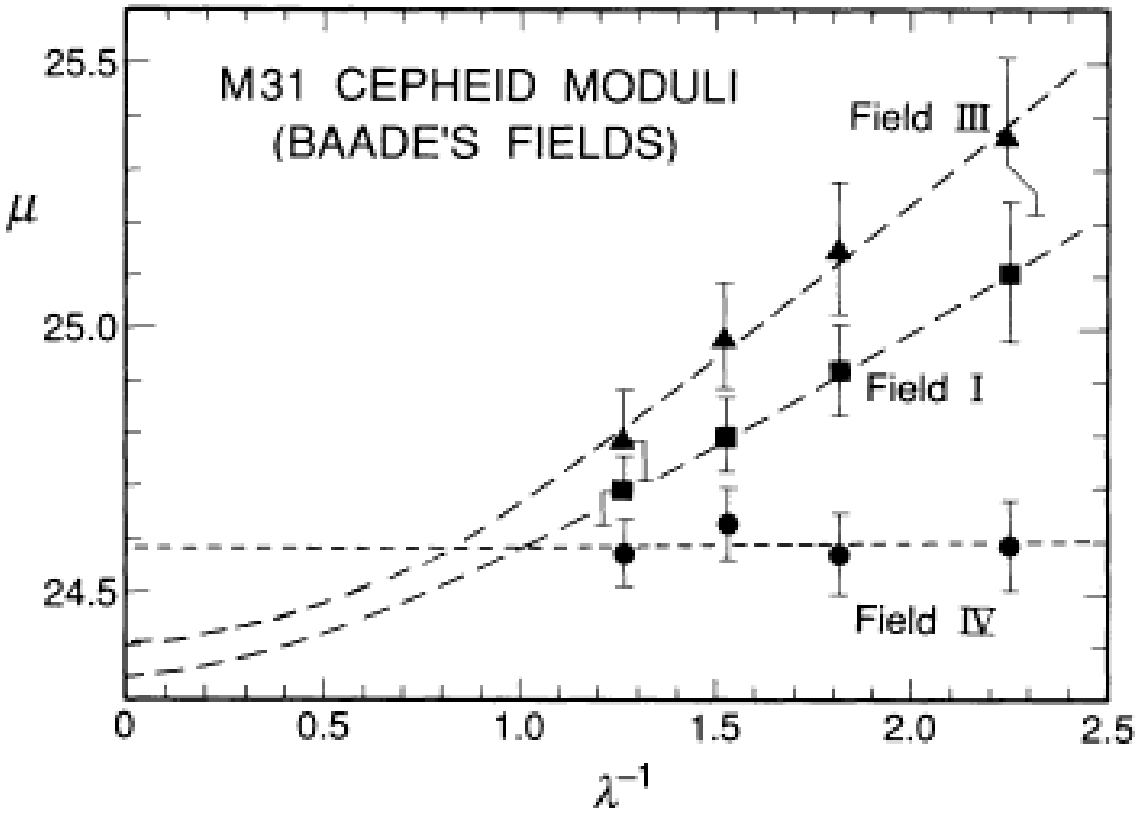


FIG. 4.—The *BVRI* apparent distance moduli for Cepheids in Baade's Fields I, III, and IV in M31 plotted against the inverse wavelength of the bandpass. The broken lines are Galactic extinction curves fitted to the apparent moduli by χ^2 minimization. There is significantly different mean extinction from field to field, (with the 10 kpc Field III being most affected, and the 20 kpc Field IV being least affected); however, the extrapolation to zero extinction at infinite wavelength indicates remarkably little residual dispersion in true modulus.

The Cepheid period luminosity relation probably depends a bit on the metallicity of the star. Unfortunately, not only is the amplitude of this effect controversial, but the is sign as well. The *HST* Distance Scale Key Project adopts a slight metallicity dependence, but probably shouldn't have, since there is very little data supporting it.

The SN 1987A Light Echo

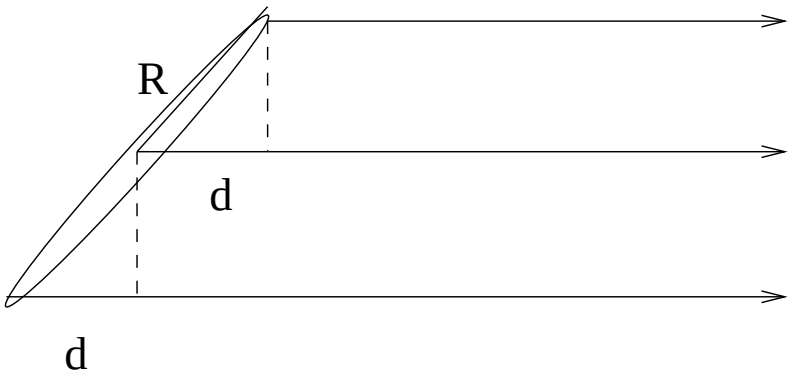
[Panagia, *et al.* 1991, *Ap.J. (Letters),* **380,** L23]
[Gould & Uza, O. 1998, *Ap.J.,* **494,** 118]

Classically, the two most reliable standard candles are Cepheids and RR Lyrae stars. Yet the Cepheid distance modulus to the Large Magellanic is ~ 18.50, while the RR Lyrae star distance is ~ 0.2 mag closer. Since the LMC is our nearest galaxy, this is disturbing. Fortunately, in 1987, a supernova went off near one of the star forming regions of the LMC. The observations of this supernovae have given us an independent distance to the galaxy.

The facts of the case are these. In February 1987, SN 1987A erupted, sending gas expanding out at $\sim 10,000$ km s^{-1}. The *Hubble Space Telescope* was not yet in orbit, but the supernovae was observed spectroscopically in the ultraviolet using the *International Ultraviolet Explorer* Telescope. Originally, *IUE* recorded only a bright continuum and very brought absorption lines from the supernova ejecta. But 83 ± 6 days after outburst, the satellilte began to see a series of narrow emission lines. The narrow lines brightened continuously until day 413 ± 24, at which time the emission plateaued. Since these lines were unresolved at *IUE* resolution, the emitting gas could not be ejecta from the supernova – it had to be from stationary surrounding material.

Several years later, *HST* imaged the supernova. In addition to seeing the (still spatially unresolved) supernova remnant, *HST* detected a ring of emission surrounding the remnant. If you assume that this ring is circular, then the axis ratio of the ring ($b/a = 0.73$) implies an inclination of $i = 43^\circ \pm 2^\circ$.

Now consider the following geometry.



Suppose the shell is circular, and that the emission lines seen by *IUE* came from the shell. The path difference between the light coming directly from the supernovae and the light echoing off the near side of the ring is

$$\Delta p = R - d = R - R\sin i = R(1 - \sin i) \tag{17.28}$$

Thus the ring's emission lines should have first appeared at time

$$t_1 - t_0 = \frac{R}{c}(1 - \sin i) \tag{17.29}$$

The ring would then continue to brighten until the light from the farthest point of the ring reached us.

$$t_2 - t_0 = \frac{R}{c}(1 + \sin i) \tag{17.30}$$

Using these two equations, the inclination of the ring would therefore be

$$\sin i = \frac{t_2 - t_1}{t_2 + t_1} \tag{17.31}$$

The physical size of the ring, R, then follows from (17.29), and the distance comes by comparing R to the observed size of the ring, $1\farcs66 \pm 0\farcs03$.

When this is done, the derived distance to SN 1987A comes out to be 51.2 ± 3.1 kpc (or $(m - M)_0 = 18.54 \pm 0.06a$), in excellent agreement with the Cepheid distance. (This distance drops to 50.1 ± 3.1 kpc if one assumes that the supernova is on the far side of the galaxy.) Moreover, the inclination derived through (17.31) is $42° \pm 5°$, almost exactly what is observed visually. This result presents extremely strong evidence that the Cepheid distance to the LMC is correct, and gets us to the nearest galaxy without the use of the distance ladder. Thus, the distance to the LMC is usually given as 50 kpc, or $(m - M) = 18.50$. This is what is used by the *HST* Distance Scale Key Project.

Note however, that the above analysis assumes that the ring is perfectly circular the recombination associated with the ionization occurred instantaneously, and the delay times determined by *IUE* are accurate. A more complicated analysis Gould & Uza which uses a re-reduction of the *IUE* data, and takes this into account delays and non-circular geometry gives a distance modulus of $(m - M)_0 < 18.37 \pm 0.04$ for a circular ring, or $(m - M) < 18.44 \pm 0.05$ for a ring with an intrinsic axis ratio of $b/a = 0.95$.

Secondary Standard Candles

Standard candles can be divided into several types. First, there are *primary* distance candles, such as parallax, moving cluster, and Baade-Wesselink measurements. These require no assumptions about their calibration. Next, there are *secondary* standard candles, that require a calibration via some primary method. These secondary methods include main-sequence fitting, Cepheids and RR Lyrae variables; without parallax, *etc.*, these would be useless.

The techniques we will now talk about are, by and large, *tertiary* standard candles. Most of them are calibrated in nearby galaxies using Cepheids and/or RR Lyrae stars. Hence the uncertainties in the primary and secondary methods will propagate into these measurements. Moreover, the uncertainties in the *testing* of tertiary standard candles can also become an issue.

There are two possible ways to test tertiary standard candles. The first, an external test, is compare the distances derived by one technique with those derived from another. If the two agree, then both methods are correct (or both are wrong in the same way). Alternatively, you can perform an internal test and derive distances to large numbers of galaxies at the same distance (say, a bunch of galaxies in a cluster). If no systematic error is seen (such as the blue galaxies being closer than the red galaxies), then you can feel somewhat confident that the method is producing good relative distances.

Luminosity Functions

It is very common in astronomy to study the number of objects present in a sample versus magnitude. Such a plot is called a luminosity function. There are several luminosity function standard candles, including planetary nebulae, globular clusters, and red giant stars. In general, these methods do not go out far enough to be used to measure the undisturbed Hubble Flow, but they are useful for measuring distances to early-type galaxies (which have no Cepheids) or checking for systematic errors on the distance ladder.

Luminosity functions are extremely useful in astronomy, not only for distance measurements, but for a whole host of other problems, from the the distribution of Milky Way field stars to the evolution of galaxies in the universe. Thus, before addressing the individual techniques, we should first consider them in a general sense.

There are two ways to think of a luminosity function. The conventional way is just to consider it as a function, $N(m)$, where N is the number of objects within a finite bin size. (Actually, one almost always uses $\log N$, rather than N.) Under this interpretation, one chooses the bin size, and creates the luminosity function by placing each object into its appropriate bin. Note, however, that due to observational errors, the observed luminosity function will *not* be the intrinsic function.

To see this, consider an intrinsic power-law luminosity function, where each bin has 10 times the counts in the previous bin, *i.e.,* bin k has 10 counts, bin $k+1$ has 100 counts, bin $k+2$ has 1000 counts, *etc.* In the presence of observational errors, some of the counts from each bin will spill over into the adjacent bins, so bin

$k+1$ will lose a fraction of its counts to bins k and $k+2$, but also gain counts from these bins. But note: all is not symmetrical. If bin $k+1$ loses 5% of its counts to bin $k+2$, but gains 5% of bin $k+2$'s counts, then bin $k+1$ is a big winner. We will therefore overestimate its numbers. The observed luminosity function is a *convolution* of the true function with the photometric error function. In general, the observed function will always be smoother and flatter than the true function.

The Eddington Correction

[Eddington 1913, *MNRAS,* **73,** 359]

There's a very tricky way to handle observational errors, which was first used by Eddington in 1913, and which is sometimes called the Eddington correction. (By far the best place to read about it is the classic 1953 book *Statistical Astronomy,* by Trumpler and Weaver.) Let $N_{obs}(m_0)$ be the observed number of counts in a bin centered at m_0, and $N_t(m)$ be the true luminosity function. To save on notation, let $x = m - m_0$, and $G(x)$ be the function which describes how the true number of counts in the bin centered at m_0 is distributed. For simplicity, let's also assume that $G(x)$ is symmetrical about zero, and, of course, normalizes to one, since no objects are lost. (Usually, you can consider G to be a Gaussian, with some dispersion, σ). With this notation, N_{obs} is just the convolution of N_t with $G(x)$,

$$N_{obs} = N_t \circ G \tag{18.01}$$

and

$$N_{obs}(m_0) = \int_{-\infty}^{\infty} N_t(m_0 - x)\, G(x)\, dx \tag{18.02}$$

Now, the key to solving this equation is to expand N_t about m_0 using a Taylor series

$$N_{obs}(m_0) = \int_{-\infty}^{\infty} \left\{ N_t(m_0) - \frac{N_t'(m_0)}{1!}x + \frac{N_t''(m_0)}{2!}x^2 + \ldots \right\}$$

$$G(x)dx \tag{18.03}$$

or

$$N_{obs}(m_0) = N_t(m_0) \int_{-\infty}^{\infty} G(x)\, dx - \frac{N_t'(m_0)}{1!} \int_{-\infty}^{\infty} x\, G(x)\, dx +$$

$$\frac{N_t''(m_0)}{2!}\int_{-\infty}^{\infty} x^2 G(x)\,dx + \ldots \tag{18.04}$$

Note that integrals of the form

$$\mu_n = \int_{-\infty}^{\infty} x^n G(x) dx \tag{18.05}$$

give the *moments* of G. So (18.05) can be re-written as

$$N_{obs}(m_0) = N_t(m_0) - \frac{\mu_1}{1!}N_t'(m_0) + \frac{\mu_2}{2!}N_t''(m_0) - \ldots \tag{18.06}$$

This is a perfectly good equation for $N_{obs}(m_0)$, except that you don't know $N_t(m_0)$, or its derivatives. But it is possible to be tricky, and say that for some set of variables, A, the following must be true

$$N_t(m_0) = N_{obs}(m_0) + A_1 N_{obs}'(m_0) + A_2 N_{obs}''(m_0) + \ldots \tag{18.07}$$

Now substitute for $N_t(m_0)$ in (18.06) using (18.07), and use (18.05) to find expressions for $N_{obs}'(m_0)$, $N_{obs}''(m_0)$, *etc.* In order for $N_{obs}(m_0)$ to equal $N_{obs}(m_0)$, each term must equal its corresponding term. So

$$\begin{aligned} A_1 - \frac{\mu_1}{1!} &= 0 \\ A_2 + \frac{\mu_2}{2!} &= 0 \\ A_3 - \frac{\mu_3}{3!} &= 0 \\ A_4 + A_2\frac{\mu_2}{2!} + \frac{\mu_4}{4!} &= 0 \end{aligned} \tag{18.08}$$

This gives us the values of A. Finally, there is one last simplification. If $G(x)$ is symmetrical, all the odd numbered terms go to zero. (You can see this easily if you graph what's going on). So

$$N_t(m_0) = N_{obs}(m_0) - \frac{\mu_2}{2!}N''_{obs}(m_0) + \left\{ \left(\frac{\mu_2}{2!}\right)^2 - \frac{\mu_4}{4!} \right\} N^{iv}_{obs}(m_0) + \ldots \tag{18.09}$$

And note: if $G(x)$ is a Gaussian function, $\mu_n = \sigma^n$. So, if we just keep the first term

$$N_t(m_0) = N_{obs}(m_0) - \frac{\sigma^2}{2!}N''_{obs}(m_0) \tag{18.10}$$

Therefore, if you know your observational error, σ, and have some idea of the shape of the luminosity function, you can easily correct your data and arrive at the true function.

Luminosity Functions as Probability Distributions

A second way of looking at a luminosity function is to consider it as a probability distribution. Let $\phi(M)$ be the intrinsic luminosity function of a set of objects, convolved with the photometric errors. (You may not know what this luminosity function is, but that's OK for now.) Since you're treating the luminosity function as a probability distribution,

$$\int_{-\infty}^{\infty} \phi(M) dM = 1 \tag{18.11}$$

Now let T be the total number of objects you are observing, and consider: how many objects are you likely to observe in a bin of size dM? Your expectation value is simply proportional to the probability amplitude at that bin, *i.e.,*

$$\lambda(M) dM = T \phi(M) dM \tag{18.12}$$

(So, if you're near the peak of a Gaussian, you'll observe more objects in a magnitude bin than you would on the tail.) Of course, the actual number of objects you will observe won't be $\lambda(M)dM$. Sometimes you'll see more, sometimes less. Since the scatter should be Poissonian about the expectation value, the probability of observing n objects in a bin is

$$p(n) = \frac{(\lambda(M)dM)^n \, e^{-\lambda(M)dM}}{n!} \tag{18.13}$$

But that's just for one bin. The total probability of finding n_1 objects in bin 1, n_2 objects in bin 2, *etc.* out to bin N is

$$P = \prod_{i=1}^{N} p_i(n) \tag{18.14}$$

This is a generalized likelihood: it tells you the probability of observing a given distribution of objects.

Taken alone, (18.14) doesn't seem too useful. After all, the probability of finding any particular configuration of objects is rather small. However, one can then ask the question, "which is more likely, luminosity function ϕ_1 or ϕ_2?" That is easily answered (*i.e.*, , ϕ_1 is twice as likely as ϕ_2). Thus, (18.14) can be thought of as being an equation for relative likelihood.

If one has thousands of objects, then (18.14) can be used directly to compute relatively likelihoods between many different models of $\phi(M)$. But the real power of the analysis occurs if the number of objects is small. To see this, let $dM \to 0$. With a infinitessimal bin size, each bin will have either zero or one object in it. We can therefore re-write (18.14) as

$$P = \prod \left\{ \frac{(\lambda(M)dM)^0 \, e^{-\lambda(M)dM}}{0!} \right\}_0 \left\{ \frac{(\lambda(M)dM)^1 \, e^{-\lambda(M)dM}}{1!} \right\}_1 \tag{18.15}$$

where the first term represents all those bins with zero objects, and the second gives all the bins with one object. Now substitute using (18.12), factor out the exponential,

$$P = \prod \left\{ e^{-T\phi(M)dM} \right\}_{\text{all bins}} \cdot \prod_{i=0}^{N} \{T\phi(M_i)dM\}_{n=1} \tag{18.16}$$

and note that since $dM \to 0$, the first term is really an integral, going from the brightest objects, down to the faintest object you can see, M_l. So

$$P = \exp\left\{ -\int_{M_l}^{-\infty} -T\phi(M)dM \right\} \prod_{i=1}^{N} T\phi(M_i)dM \tag{18.17}$$

The first term in the equation is just a normalization; the second term is just the product of the individual probabilities.

Let's now apply this to a specific example: determining distance. We can translate absolute magnitude to apparent magnitude via $M = m - \mu$, so

$$P = \exp\left\{ - \int_{m_l - \mu}^{-\infty} -T\phi(M)dM \right\} \prod_{i=1}^{N} \phi(m_i - \mu)dm \qquad (18.18)$$

What is the most likely distance modulus, μ? It's the value that produces the highest relatively likelihood, P.

Three additional notes. First, from a computaional point of view, it is usually easier to work in terms of log probabilities. In that case, the products in the above equations become summations, *i.e.,*

$$\ln P = - \int_{m_l - \mu}^{-\infty} -T\phi(M)dM + N \ln T + \sum_{i=i}^{N} \ln \phi(m_i - \mu) \quad (18.19)$$

(where we've dropped the $N \ln dm$ term, since it's just a constant). Second, if one were to investigate all possible distance moduli (and all possible values of T), then you can re-normalize the likelihoods to sum to one, since the solution must be in there somewhere. Finally, this type of analysis doesn't tell you whether any of the fits are good or bad; it just let's you know which are best.

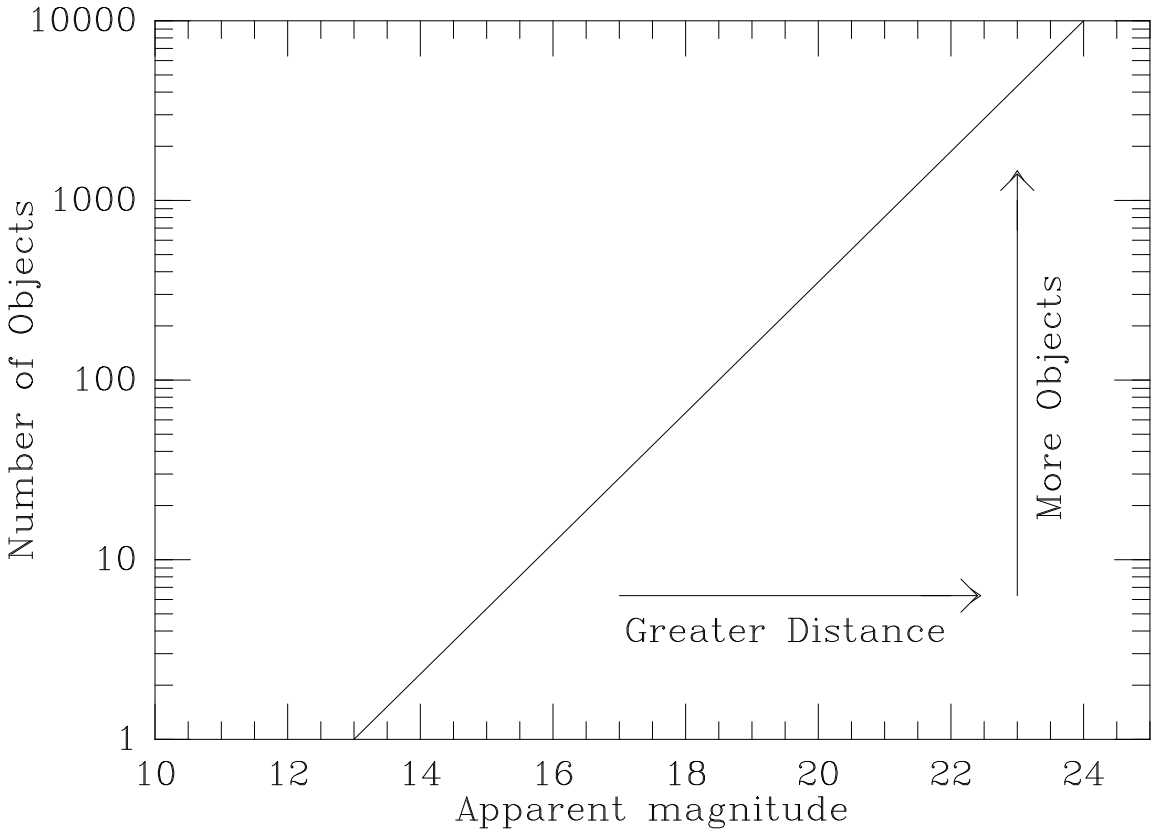
Some problems in astronomy are simple when you think of luminosity functions as probability functions, and impossible if you work with them as luminosity functions. (For instance, how do you bin a total of 5 objects?) And the reverse is also true. So both tools are needed.

Power-law Luminosity Functions

Consider a luminosity function which is a power law, *i.e.*, $N(\mathcal{L}) \propto \mathcal{L}^{\alpha}$, or

$$N(m) = am + b \tag{18.20}$$

where a and b are two constants. In general, of course, these constants are not known ahead of time. (In fact, they are generally what you are looking for.) Now, let's observe two galaxies that look identical on the sky, but let one be 10 times further away than the other. Obviously, the objects in the more distant galaxy will five magnitudes fainter than that of the nearer galaxy. But that more distant galaxy (in order to have the same apparent magnitude) would have to be 100 times more luminous. Therefore, the number of objects at each magnitude would be 100 times greater. Consequently, the $\log N$ *vs.* m diagram for both galaxies would be the same! This is a general rule – you can't tell anything from a power-law luminosity distribution.

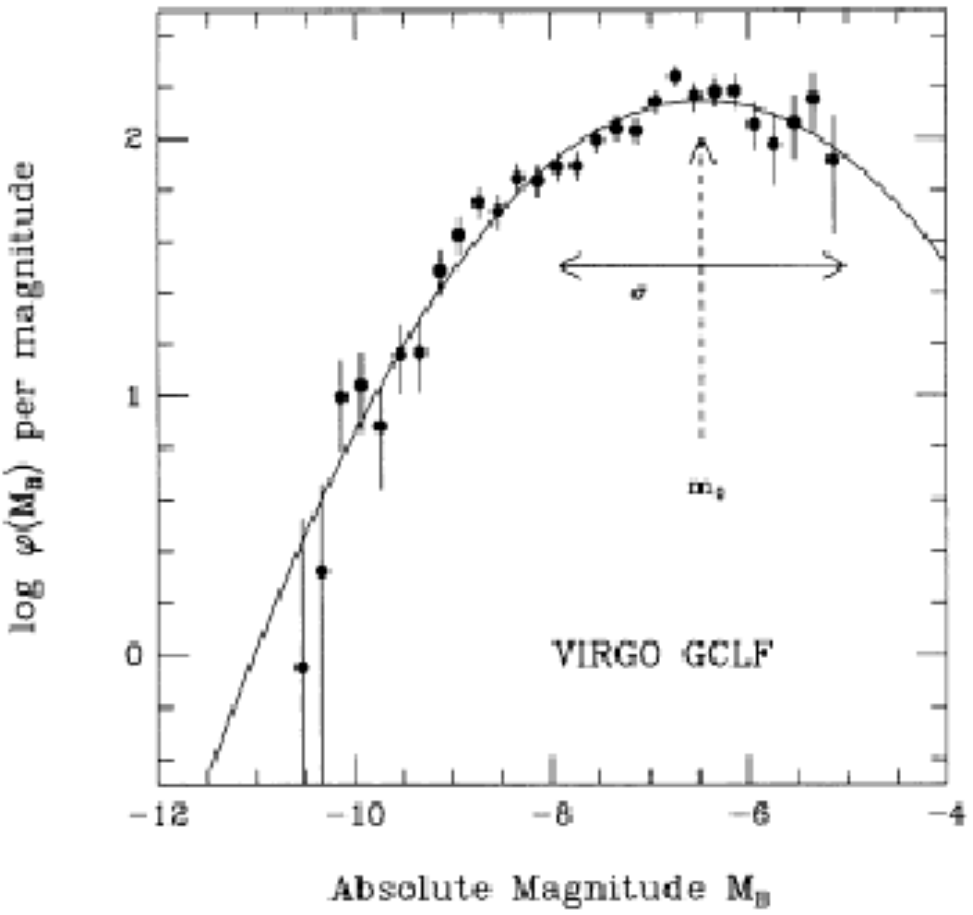


The Globular Cluster Luminosity Function

When put at the distance of $\gtrsim 1$ Mpc away, globular clusters look like a faint stars. But some galaxies have alot of them, making them easy to count, even with background contamination. (This is especially true in smooth, elliptical and S0 systems.) In most galaxies, the GCLF is a log Gaussian, *i.e.,* when plotted as $\log N$ *vs.* m, it is a Gaussian,

$$\log N(m) \propto e^{(M-M_0)^2/2\sigma^2} \tag{18.21}$$

with $M_0 \sim -7.3$ in the V-band, and $\sigma \approx 1.2$. Thus by observing the apparent magnitude of the GCLF peak, m_0, it's possible to estimate distance.



The observed globular cluster luminosity function for M87, and the best-fit Gaussian.

Tip of the Red Giant Branch

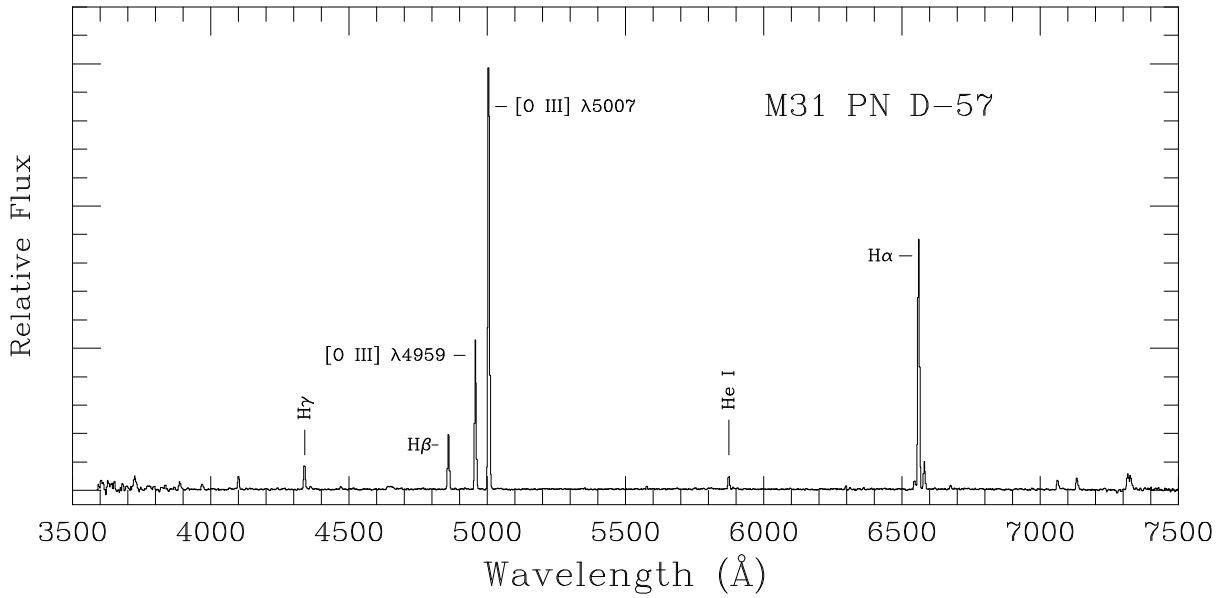
According to stellar evolution, when a star runs out of hydrogen in its core, it develops an inert helium core and evolves to the red in the HR diagram. Once on the giant branch, the stars increase their brightness, until helium itself fuses, via the triple-α process. Due to the physics of stellar interiors, this fusion takes place at about the same absolute luminosity for all stars with initial masses less than $\sim 2.5\mathcal{M}_{\odot}$. Thus, the tip of the red giant branch (TRGB) can be a useful standard candle.

Detecting individual red giants in distant galaxies is difficult; measuring their brightness is even harder. (Galaxies have many, many, many red giants, and their fluxes blur together.) In addition, although the total luminosity of the red giant stars do not depend on their mass or metallicity, the flux in a particular filter does. (For example, metal-poor TRGB stars are more luminous in the I band than metal-rich TRGB stars.) Nevertheless, by finding the apparent magnitude of the RGB tip, and comparing it to the absolute magnitude of the TRGB in galactic globular clusters, distances can be estimated.

The Planetary Nebula Luminosity Function

[Jacoby 1989, *Ap.J.*, **339,** 39]
[Ciardullo *et al.* 1989, *Ap.J.*, **339,** 53]
[Ciardullo *et al.* 2002, *Ap.J.*, **577,** 31]

When seen in other galaxies, planetary nebulae are point sources. The central stars of bright PN are stars that have just pealed off the asymptotic giant branch, and are moving across the HR diagram (to the blue), and about to stop producing energy and become white dwarfs. In the meantime, however, they are extremely bright (*i.e.,* just as luminous as they were on the tip of the AGB) and extremely hot. Most of their energy is trapped by their surrounding nebulae, and goes into ionizing hydrogen. This energy is released in a series of emission lines, the brightest of which (usually by far) is [O III] λ5007.



When you observe a PN in a distant galaxy, you will see light at 5007 Å, but no light at adjacent wavelengths. (This is in contrast to stars which are continuum sources.) Therefore, they are easy to identify.

The premise behind the planetary nebula luminosity function is that the luminosity function of PN in [O III] λ5007 is always the same

$$N(m) \propto e^{0.307M} \left\{ 1 - e^{3(M^* - M)} \right\} \tag{18.22}$$

where M^* is the magnitude of the brightest possible PN. Since observations of PN are not done in any "standard" broadband system, the conversion between PN flux and PN magnitude is rather arbitrary (*i.e.,* , was invented on the spare of the moment one Thursday morning in 1988)

$$m_{5007} = -2.5 \log F_{5007} - 13.74 \tag{18.23}$$

where F_{5007} is the flux contained in the [O III] λ5007 emission line in ergs cm^{-2} s^{-1}.

The PNLF technique has been tested rather thoroughly, both against other "reliable" techniques (such as Cepheid distances), and using multiple galaxies within a cluster. The results indicate the PNLF does not depend on the color, metallicity, or age of the galaxy, and that, PNLF measurements are good to $\sim 10\%$ or better. (The lack of dependence on metallicity is understandable; since O^{++} is the major coolant in the nebula, decreasing its abundance only serves to drive up the electron temperature and increase the emission per ion. The invariance with population age is more of a mystery, and probably has something to do with blue straggler stars.)

The PNLF is the only commonly used technique that works equally well in both spiral and elliptical galaxies. It is therefore extremely for cross-checking the results of different methods, and looking for systematic errors in the distance ladder.

Surface Brightness Fluctuations

[Tonry & Schneider 1988, *A.J.,* **96,** 807]
[Tonry *et al.* 2001, *Ap.J.,* **546,** 681]
[Jensen *et al.* 2003, *Ap.J.,* **583,** 712]

Like the PNLF method, the surface brightness fluctuation (SBF) technique was also invented in 1988. Unlike the PNLF, it has the capability of reaching galaxies in the unperturbed Hubble flow.

Consider the isophotes (contour of equal surface brightness) of a smooth elliptical or S0 galaxy. Let n be the surface density of stars along the isophote, $\overline{\mathcal{L}}$ be the "mean" stellar luminosity, and r be the distance to the galaxy. The mean flux recorded in each pixel of your detector is, of course, independent of distance: as you move a galaxy further away, the flux from each star declines as $1/r^2$, but the number of stars per unit area on the sky increases as r^2. So the mean flux (number of stars times flux from each star) is

$$\langle F \rangle \propto (nr^2)(\overline{\mathcal{L}}/r^2) \propto n\overline{\mathcal{L}} \tag{18.24}$$

In reality, the number of stars in each pixel won't be exactly the same – there will be some statistical fluctuation in the stellar density due to Poisson statistics. (In other words, if the average number of stars per pixel is 100, the expected pixel-to-pixel fluctuation will be $\sqrt{100}$, or 10%.) Consequently, the pixel-to-pixel flux will have a standard deviation of

$$\sigma_{\langle F \rangle} \propto (nr^2)^{1/2}(\overline{\mathcal{L}}/r^2) \tag{18.25}$$

a variance of

$$\sigma^2_{\langle F \rangle} \propto (nr^2)(\overline{\mathcal{L}}/r^2)^2 \propto n\overline{\mathcal{L}}^2/r^2 \tag{18.26}$$

and a fractional variance of

$$\frac{\sigma^2_{\langle F \rangle}}{\langle F \rangle} = \frac{n\overline{\mathcal{L}}^2}{\overline{\mathcal{L}}}\frac{1}{r^2} = n\overline{\mathcal{L}}\frac{1}{r^2} \tag{18.27}$$

(Note that this, in fact, defines what "mean luminosity" means)

$$\overline{\mathcal{L}} = \frac{\sum_i N_i \mathcal{L}_i^2}{\sum_i N_i \mathcal{L}_i} \tag{18.28}$$

Thus, by measuring the fractional pixel-to-pixel "noise", it is possible to estimate the distance, r.

Of course, the method has its difficulties. Pixel-to-pixel noise comes from a variety of sources, including read-noise of the detector, shot-noise from photon statistics, noise from small knots of emission and/or dust, and noise from faint point sources (such as foreground stars, background galaxies, and the galaxy's globular clusters). So each of these must be modeled and removed.

In addition, the method assumes that $\overline{\mathcal{L}}$ is the same for every galaxy. It is not. The value of $\overline{\mathcal{L}}$ (or equivalently $\overline{M}$) is fainter in redder galaxies than it is in bluer galaxies. (It is also bandpass dependent!) The exact calibration of the SBF fluctuation magnitude has gone through many iterations (particularly in the I-band), and was given by Tonry *et al.* (2001) as

$$\overline{M} = -1.74 + 4.5\,[(V-I)_0 - 1.15] \tag{18.29}$$

where $1.0 < (V-I)_0 < 1.3$ is the intrinsic color of the galaxy and $\overline{M_I}$ is the surface brightness fluctuation as measured in the I-band. (Galaxies don't get much redder than this, while bluer galaxies generally have dust and star formation which makes the measurement of surface brightness fluctuations impossible. So $\overline{M_I}$ is calibrated only over a finite range.)

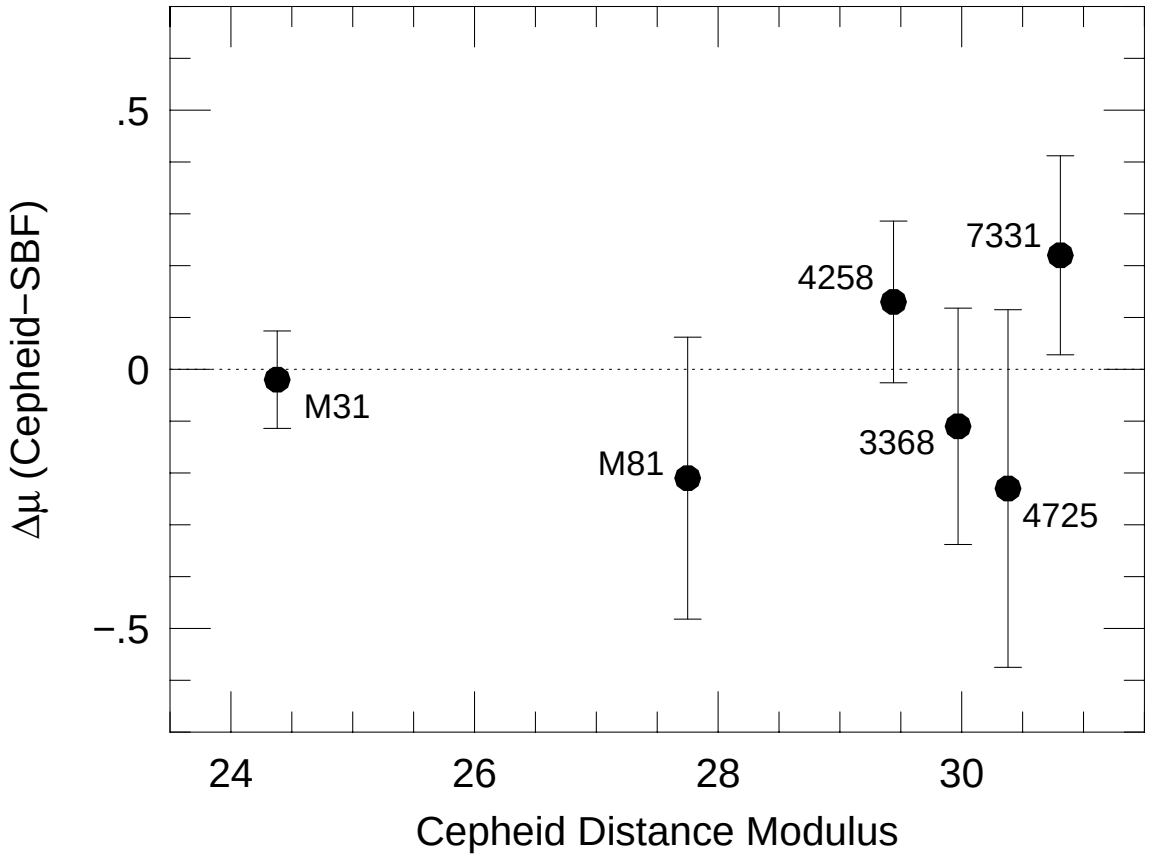
In theory, surface brightness fluctuations can be measured in any filter bandpass (and it may be best to measure it in the infrared, where $\overline{M}$ is brightest), but the largest set of measurements have been done in I.

The Results from Surface Brightness Fluctuations

Over 300 galaxies have SBF distance measurements. Since many of these galaxies are distant enough to be in the unperturbed Hubble Flow, we should be able to obtain a Hubble Constant from its measurements. Unfortunately, doing so demonstrates the potential problems with Hubble Constant measurements.

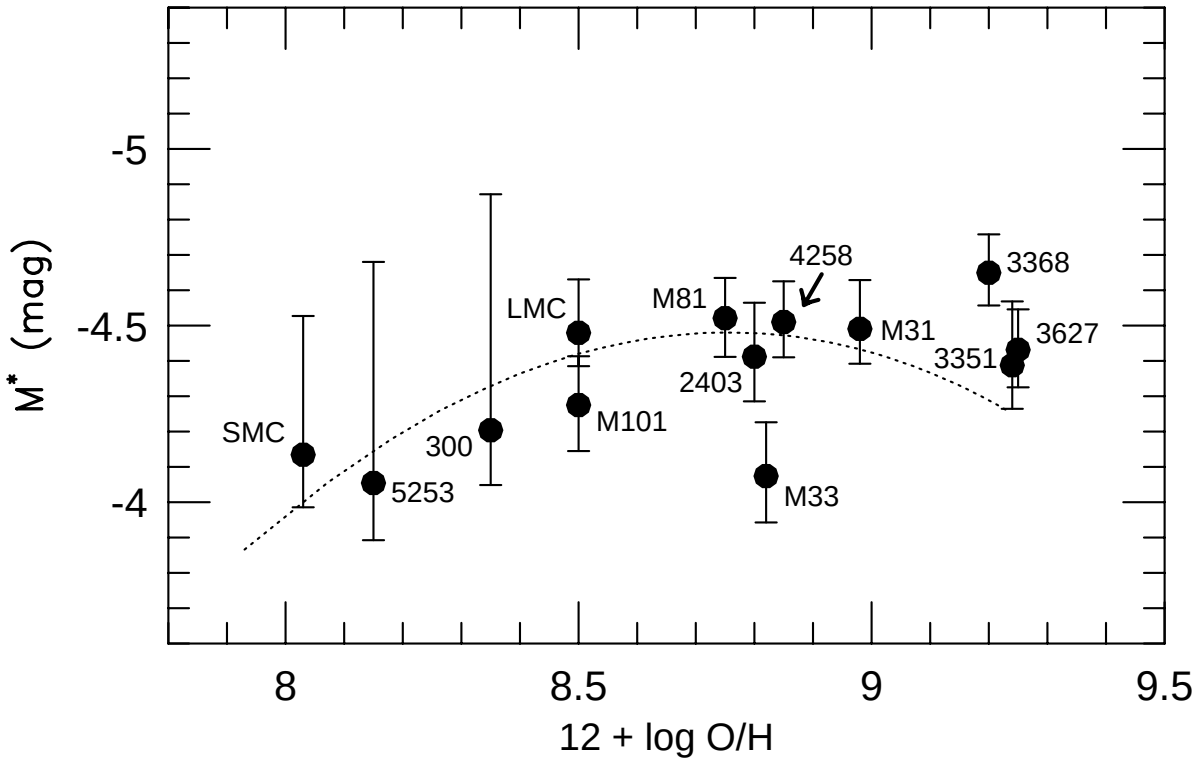
The SBF method is a tertiary standard candle, so it must be calibrated with another technique. The most direct way to do this is through pulsating stars. RR Lyr variables aren't suitable since they're bright enough to be seen outside the Local Group. That leaves Cepheid variables. Unfortunately, Cepheids are Pop I objects, while SBF measurements can only be made in smooth galaxies without any dust (*i.e.,* ellipticals and S0's). Thus, in order to calibrate the SBF measurements with Cepheids, either use a) different galaxies in a group (and assume they're all at the same distance, or b) spiral galaxies with sizeable bulges.

There are 6 galaxies that satisfy the latter condition; if you calibrate the SBF with these galaxies, you get $H_0 = 70 \pm 5$ km/s/Mpc (with an additional zero point uncertainty of ± 6). This is the calibration that was used by the *HST* Key Project. However, a revised calibration based on infrared SBF measurements yield a value closer to $H_0 = 80 \pm 5$ km/s/Mpc.

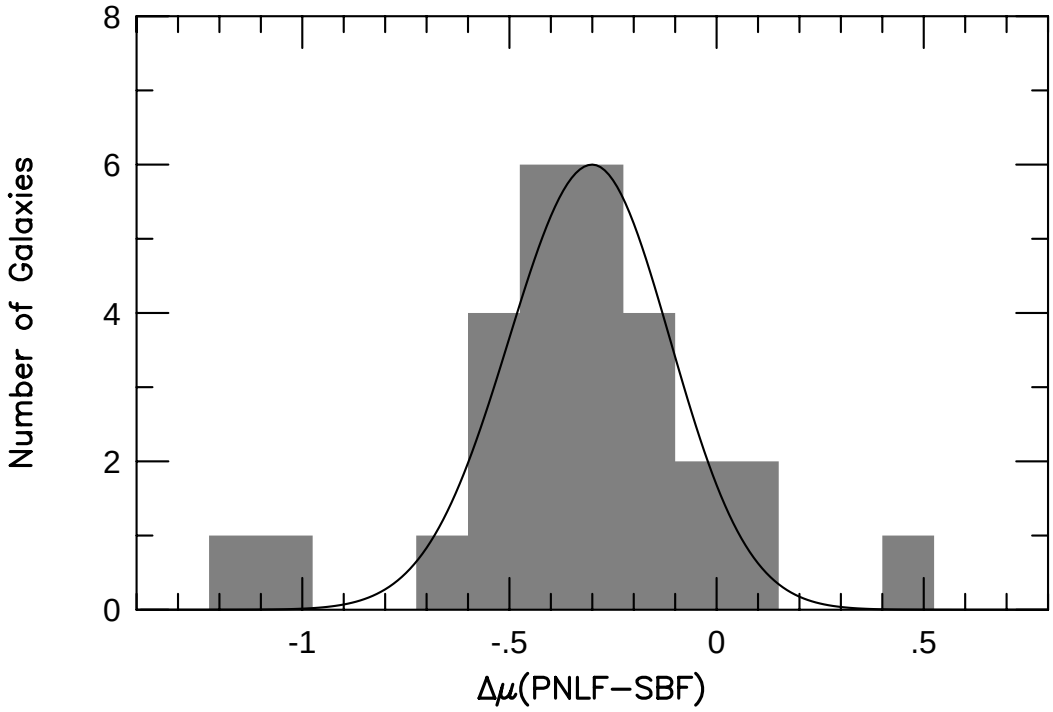


The Cepheid calibration of the SBF method. The plot shows the residuals between the mean calibration and the distance derived for each galaxy.

However, SBF distances can also be calibrated indirectly via the PNLF (*i.e.,* Cepheid → PNLF → SBF). Since planetary nebulae are found in all galaxies, there is substantial overlap in the techniques. To date, 13 galaxies have both PNLF and Cepheid measurements, and 28 galaxies have PNLF and SBF data. If you compute H_0 in this way, then H_0 is 15% larger. Much of this offset has now been explained – the zero point of the SBF method changed yet again.



The Cepheid calibration of the PNLF cutoff magnitude, M^*. The galaxies' Cepheid distances are used to translate the apparent magnitude of the cutoff, m^*, to absolute magnitude.



The difference between the PNLF-derived distance modulus and SBF-derived distance modulus for 28 galaxies measured with both techniques. The scatter is consistent with the expected errors (the curve), but the mean is quite different – 0.3 mag, or 15%.